BLOCK KRYLOV SUBSPACE EXACT TIME INTEGRATION OF LINEAR ODE SYSTEMS PART 1: ALGORITHM DESCRIPTION*

MIKE A. BOTCHEV†

Abstract. We propose a time-exact Krylov-subspace-based method for solving linear ODE (ordinary differential equation) systems of the form $y' = -Ay + g(t)$, where $y(t)$ is the unknown function. The method consists of two stages. The first stage is an accurate polynomial approximation of the source term $g(t)$, constructed with the help of the truncated SVD (singular value decomposition). The second stage is a special residual-based block Krylov subspace method.

The accuracy of the method is only restricted by the accuracy of the polynomial approximation and by the error of the block Krylov process. Since both errors can, in principle, be made arbitrarily small, this yields, at some costs, a time-exact method.

Key words. Krylov subspace; matrix function; block Arnoldi process; block Lanczos process; exponential time integrators; matrix exponential residual; truncated SVD; proper orthogonal decomposition

AMS subject classifications. 65F60, 65F10, 65F30, 65N22, 65L05

1. Problem formulation. Consider initial-value problem (IVP)

$$\begin{cases} y' = -Ay + g(t), & \\ y(0) = v, & t \in [0, T], \end{cases} \tag{1.1}$$

where $y(t)$ is the unknown vector function, $y : \mathbb{R} \to \mathbb{R}^n$, and the matrix $A \in \mathbb{R}^{n\times n}$, vector function $g : \mathbb{R} \to \mathbb{R}^n$, and vector $v \in \mathbb{R}^n$ are given.

Let $\tilde{y}(t) \equiv y(t) - v$ (meaning that $\tilde{y}(t) = y(t) - v$ for all t). Note that the function $\tilde{y}(t)$ satisfies IVP

$$\begin{cases} \tilde{y}' = -A\tilde{y} + \tilde{g}(t), & \\ \tilde{y}(0) = 0, & t \in [0, T], \end{cases} \tag{1.2}$$

where $\tilde{g}(t) \equiv g(t) - Av$. We will assume that the IVP (1.1) is brought to the equivalent form (1.2) and, for simplicity, we omit the tilde sign $\tilde{}$ in (1.2).

2. Polynomial approximation. We now describe the first stage of the method, the best fit polynomial approximation of the source term $g(t)$. Choose s points $0 = t_1 < t_2 < \cdots < t_{s-1} < t_s = T$ on the time interval $[0, T]$. The polynomial approximation is based on the truncated SVD (singular value decomposition) of the matrix

$$\tilde{G} = \begin{bmatrix} g(t_1) & g(t_2) & \dots & g(t_s) \end{bmatrix} \in \mathbb{R}^{n\times s},$$

whose columns are samples $g(t_i)$, $i = 1, \dots s$, of the vector function $g(t)$. More precisely, let

$$\tilde{G} = \tilde{U}\tilde{\Sigma}\tilde{V}^T, \quad \tilde{\Sigma} = \mathrm{diag}(\sigma_1, \dots, \sigma_s) \in \mathbb{R}^{s\times s}, \quad \sigma_1 \geqslant \dots \geqslant \sigma_s \geqslant 0, \tag{2.1}$$

*This manuscript is created on September 23, 2011.

†Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, the Netherlands, `mbotchev@na-net.ornl.gov`.

be the thin SVD [3, Section 2.5.4], where the matrices $\tilde{U} \in \mathbb{R}^{n\times s}$ and $\tilde{V} \in \mathbb{R}^{s\times s}$ have orthonormal columns $u_1, \dots, u_s$ and $v_1, \dots, v_s$, respectively. An approximation to $\tilde{G}$ can be obtained by truncating the SVD as

$$\tilde{G} = \tilde{U}\tilde{\Sigma}\tilde{V}^T = \sum_{i=1}^{s} \sigma_i u_i v_i^T \approx \sum_{i=1}^{m} \sigma_i u_i v_i^T = U\Sigma V^T, \quad m < s, \tag{2.2}$$

where $\Sigma \in \mathbb{R}^{m\times m} = \mathrm{diag}(\sigma_1, \dots, \sigma_m)$ and the matrices $U \in \mathbb{R}^{n\times m}$ and $V \in \mathbb{R}^{s\times m}$ are formed by the first m columns of $\tilde{U}$ and $\tilde{V}$, respectively. Denote the obtained approximate matrix by $G = U\Sigma V^T$. If follows from (2.2) that the SVD of $\tilde{G} - G$ is readily available as $\sum_{i=m+1}^{s} \sigma_i u_i v_i^T$. Hence, for the 2-norm and Frobenius norm of the error $\tilde{G} - G$ holds [3, Section 2.5.3]:

$$\|\tilde{G} - G\|_2 = \sigma_{m+1}, \quad \|\tilde{G} - G\|_F^2 = \sigma_{m+1}^2 + \dots + \sigma_s^2.$$

Looking at SVD identity (2.1) columnwise, we see that every sample value $g(t_i)$ of the function $g(t)$ can be approximated by a linear combination of the vectors $u_1, \dots, u_m$:

$$\begin{aligned} g(t_i) &= (\sigma_1 v_{i1})u_1 + (\sigma_2 v_{i2})u_2 + \dots + (\sigma_s v_{is})u_s \\ &\approx (\sigma_1 v_{i1})u_1 + (\sigma_2 v_{i2})u_2 + \dots + (\sigma_s v_{im})u_m, \end{aligned}$$

where v_{ij} are the entries of the unitary matrix V. Following the approach of [2], we consider the coefficients of these linear combinations, namely $\sigma_j v_{ij}$, $j = 1, \dots, m$, as values of some unknown functions $f_j(t)$ at t_i. These functions can be easily approximated, at a low cost (typically $m \ll n$) and with a very high accuracy, by a polynomial fit [2]. This yields

$$\begin{aligned} g(t_i) &\approx f_1(t_i)u_1 + f_2(t_i)u_2 + \dots + f_m(t_i)u_m \\ &\approx p_1(t_i)u_1 + p_2(t_i)u_2 + \dots + p_m(t_i)u_m. \end{aligned} \tag{2.3}$$

For simplicity, we assume that all the best-fit polynomials have the same the order r. Packing the polynomials $p_j(t)$, $j = 1, \dots, m$, in one polynomial vector function $p(t) = (p_1(t), \dots, p_m(t))^T$, we obtain a polynomial approximation

$$g(t) \approx Up(t). \tag{2.4}$$

There are three sources contributing to the approximation error here. First, the quality of the approximation is influenced by the choice of the sample points $t_1, \dots, t_s$. Second, by the number of terms m in the SVD truncation (2.2) and, finally, by the polynomial best fit in (2.3). All these errors can be easily controlled when the approximation is constructed [2], thus giving possibility for an adaptive approximation procedure. With (2.4), the original initial-value problem (1.2) takes the form

$$\begin{cases} y' = -Ay + Up(t), \\ y(0) = 0, \qquad\qquad t \in [0, T], \end{cases} \tag{2.5}$$

We now introduce a block Krylov subspace method to solve this problem.

3. Residual-based block Krylov subspace method. To construct a Krylov subspace block iterative method for solving (2.5), we use the exponential residual concept described in [1]. Choosing the initial guess $y_0(t)$ to be a zero vector function, we see that the corresponding initial residual is

$$r_0(t) = -Ay_0(t) - y_0'(t) + Up(t) = Up(t). \tag{3.1}$$

We follow the approach of [1], where the approximate solution $y_k(t)$ at Krylov iteration k is obtained as

$$y_k(t) = y_0(t) + \xi_k(t).$$

Here the vector function $\xi_k(t)$ is the Krylov subspace approximate solution of the correction problem

$$\begin{cases} \xi' = -A\xi + r_0(t), & \\ \xi(0) = 0, & t \in [0, T], \end{cases} \tag{3.2}$$

Note that if $\xi_k(t)$ solves (3.2) exactly then $y_k(t)$ is the sought-after exact solution of (2.5). We solve (3.2) by projecting it onto a block Krylov subspace defined as

$$\mathcal{K}_k(A, U) \equiv \operatorname{span}\left\{U, AU, A^2U, \dots, A^{k-1}U\right\},$$

with dimension at most $k{\cdot}m$. An orthonormal basis for this subspace can be generated by the block Arnoldi or Lanczos process described e.g. in [5, 4]. The process produces, after k block steps, matrices

$$V_{[k+1]} = \begin{bmatrix} V_1 & V_2 & \dots & V_{k+1} \end{bmatrix} \in \mathbb{R}^{n\times(k+1)m}, \qquad H_{[k+1,k]} \in \mathbb{R}^{(k+1)m\times km}.$$

Here $V_i \in \mathbb{R}^{n\times m}$, V_1 is taken to be the matrix U produced by the truncated SVD (2.2) and $V_{[k+1]}$ has orthonormal columns spanning the Krylov subspace, namely,

$$\operatorname{colspan}(V_{[k]}) = \mathcal{K}_k(A, U).$$

The matrix $H_{[k+1,k]}$ is block upper Hessenberg, with $m{\times}m$ blocks H_{ij}, $i = 1, \dots, k{+}1$, $j = 1, \dots, k$. The matrices $V_{[k+1]}$ and $H_{[k+1,k]}$ satisfy the block Arnoldi (Lanczos) decomposition [5, 4]

$$AV_{[k]} = V_{[k+1]}H_{[k+1,k]} = V_{[k]}H_{[k,k]} + V_{k+1}H_{k+1,k}E_k^T, \tag{3.3}$$

where $H_{k+1,k}$ is the only nonzero block in the last $k+1$ block row of $H_{[k+1,k]}$ and $E_k \in \mathbb{R}^{n\times k}$ is formed by the last m columns of the $km \times km$ identity matrix.

The Krylov subspace solution $\xi_k(t)$ is computed as

$$\xi_k(t) = V_{[k]}u(t),$$

where $u(t)$ solves the projected IVP

$$\begin{cases} u'(t) = -H_{[k,k]}u(t) + V_{[k]}^T r_0(t), & \\ u(0) = 0, & t \in [0, T]. \end{cases} \tag{3.4}$$

Note that

$$V_{[k]}^T r_0(t) = V_{[k]}^T V_1 p(t) = E_1 p(t),$$

where $E_1 \in \mathbb{R}^{km\times m}$ is formed by the first m columns of the $km \times km$ identity matrix. Using (3.1), (3.3) and (3.4), we see that for the exponential residual $r_k(t)$ of the solution $y_k(t)$ holds

$$\begin{aligned} r_k(t) &= -Ay_k - y_k' + Up(t) = -Ay_0 - y_0' - AV_{[k]}u(t) - V_{[k]}u'(t) + Up(t) = \\ &= r_0(t) - AV_{[k]}u(t) - V_{[k]}u'(t) = \\ &= r_0(t) - (V_{[k]}H_{[k,k]} + V_{k+1}H_{k+1,k}E_k^T)u(t) - V_{[k]}u'(t) = \\ &= r_0(t) - V_{[k]}(H_{[k,k]}u(t) + u'(t)) - V_{k+1}H_{k+1,k}E_k^T u(t) = \\ &= r_0(t) - V_k E_1 p(t) - V_{k+1}H_{k+1,k}E_k^T u(t) = -V_{k+1}H_{k+1,k}E_k^T u(t). \end{aligned} \tag{3.5}$$

A similar expression for the exponential residual is obtained in [1] for a non-block Krylov subspace method. There are two important messages relation (3.5) provides. First, the residual can be computed efficiently during the iteration process because the matrices V_{k+1} and $H_{k+1,k}$ are readily available in the Arnoldi or Lanczos process. Second, the residual after k block steps has the same form as the initial residual (3.1), namely it is a matrix of m orthonormal columns times a time dependent vector function. This allows for a restart in the block Krylov method: set $y_0(t) := y_k(t)$, then relation (3.1) holds with $U := V_{k+1}$ and $p(t) := -H_{k+1,k}E_k^T u(t)$. The just described correction with k block Krylov iterations can then be repeated, which results in a restarted block Krylov subspace method for solving (2.5).

REFERENCES

[1] M. A. Botchev. Residual, restarting and Richardson iteration for the matrix exponential. Memorandum 1928, Department of Applied Mathematics, University of Twente, Enschede, November 2010. `http://eprints.eemcs.utwente.nl/18832/`.

[2] M. A. Botchev, G. L. G. Sleijpen, and A. Sopaheluwakan. An SVD-approach to Jacobi-Davidson solution of nonlinear Helmholtz eigenvalue problems. *Lin. Algebra Appl.*, 431:427–440, 2009. `http://dx.doi.org/10.1016/j.laa.2009.03.024`.

[3] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London, third edition, 1996.

[4] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Book out of print, 2000. `www-users.cs.umn.edu/~saad/books.html`.

[5] H. A. van der Vorst. *Iterative Krylov methods for large linear systems*. Cambridge University Press, 2003.